The EMBO Journal vol.6 no.13 pp.3985–3994, 1987

The chicken progesterone receptor: sequence, expression and functional analysis

H.Gronemeyer, B.Turcotte, C.Quirin-Stricker, M.T.Bocquel, M.E.Meyer, Z.Krozowski, J.M.Jeltsch, T.Lerouge, J.M.Garnier and P.Chambon

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, Faculté de Médicine, 67085 Strasbourg Cédex, France

Communicated by P.Chambon

The complete mRNA sequence of the chicken progesterone receptor (cPR) has been determined. Expression of the cloned cDNA both *in vivo* and *in vitro* produces a protein that has the same apparent mol. wt on SDS–polyacrylamide gels as the 'natural' cPR form B (109 kd) as determined by immunoblotting and photoaffinity labelling. When expressed in HeLa or in Cos-1 cells the 'cloned' cPR displays hormone binding characteristics indistinguishable from the 'natural' receptor and, in the presence of progestins, exhibits 'tight nuclear binding'. A protein corresponding in size to the cPR form A (79 kd) could be detected by expressing *in vivo* and *in vitro* an N-terminally truncated cPR starting at methionine 128. A protein of the same apparent mol. wt results from internal initiation during *in vitro* translation. In contrast, such a protein was barely detectable after *in vivo* expression of the cPR cDNA in Cos-1 cells. These results suggest that form A is generated by an oviduct cell specific process involving either internal initiation of translation and/or proteolysis in the vicinity of methionine-128. The cPR contains two highly conserved regions C and E, a characteristic of the steroid/thyroid hormone receptor supergene family. By expression of a series of cPR deletion mutants, region E could be defined as the hormone binding domain whereas region C is indispensable for the tight nuclear association of the progestin–receptor complex. In the presence of progestins, the cloned cPR efficiently trans-activates transcription from the long terminal repeat region (LTR) of the mouse mammary tumor virus (MMTV). Deletion of the entire N-terminal region A/B or of the hormone binding domain E results in a 100-fold reduction of transcriptional activation. No stimulation of transcription can be detected when the C-terminal deletion extends into region C, indicating that this region is involved in the recognition of the hormone responsive element (HRE) of the MMTV LTR.

Key words: progesterone receptor forms A and B/functional domains/MMTV LTR steroid hormone responsive element/activation of transcription

Introduction

The chicken and human progesterone receptors are unique in the 'superfamily' of steroid/thyroid hormone receptors (for recent reviews see Green *et al.*, 1987; Gronemeyer *et al.*, 1987) in that two receptor proteins of different mol. wts have been observed at approximately equimolar ratios in the cytosol of chicken oviduct tubular gland cells and human breast cancer T-47D cells (Schrader *et al.*, 1980; Gronemeyer *et al.*, 1983; Horwitz *et al.*, 1985; Gronemeyer and Govindan, 1986). Different functions have been attributed to the two chicken oviduct progesterone receptor (cPR) forms (Schrader *et al.*, 1980) and it has been suggested that they correspond to two different proteins (Birnbaumer *et al.*, 1983). However, immuno-analysis and peptide mapping of the photoaffinity-labelled proteins have provided strong evidence that the 79 kd form A and the 109 kd form B of the cPR are structurally (Gronemeyer *et al.*, 1983) and immunologically (Gronemeyer *et al.*, 1985) closely related. These results were confirmed by others using both poly (Tuohimaa *et al.*, 1984) and monoclonal (Sullivan *et al.*, 1986) antibodies. Furthermore, in contrast to previous reports (for review, see Schrader *et al.*, 1980), we have demonstrated that not only the form A of the cPR, but also the form B could bind to DNA (Gronemeyer *et al.*, 1985). This binding was confirmed by DNase I footprinting experiments using the hormone responsive element of the chicken lysozyme gene (v.der Ahe *et al.*, 1986).

The cloning of partial cDNA sequences of the chicken progesterone receptor has been reported recently (Conneely *et al.*, 1986; Jeltsch *et al.*, 1986). We describe here the cloning and sequencing of the entire cPR cDNA, its expression *in vivo* and *in vitro* and a functional analysis of the various domains of the expressed receptor. We also present evidence indicating that the cPR form A corresponds to an N-terminally truncated form B, generated by either cell-specific proteolysis or internal initiation of translation.

Results

The chicken progesterone receptor mRNA sequence

We have previously reported (Jeltsch *et al.*, 1986) the identification of several cDNA clones of the cPR based on (i) epitope selection, (ii) protein sequence data and (iii) the presence of a cysteine-rich region, which is characteristic of the steroid/thyroid hormone receptor family (for reviews see Green and Chambon, 1986; Green *et al.*, 1987; Gronemeyer *et al.*, 1987). By cloning overlapping cDNAs (see legend to Figure 1) and the corresponding genomic sequences (to be published separately) we have determined that the cPR mRNA sequence is 4472 nucleotides long with a 366 nucleotide-long 5′- and a 1745 nucleotide-long 3′-untranslated region (Figure 1). The AUG initiation codon at position 367 which is preceded by an in-frame termination codon at position 208 (stars in Figure 1) initiates a 786 amino acid-long open reading frame. An AUG codon, directly followed by a UGA termination codon is present in the 5′-untranslated region of the cPR mRNA (positions 290–295; underlined in Figure 1), where it might influence the efficiency of translation (for discussion of this possibility, see Kozak, 1984b; Green *et al.*, 1987 and Gronemeyer *et al.*, 1987). Similar to other receptor cDNAs (for review see Gronemeyer *et al.*, 1987) the cPR cDNA contains an unusually long 3′-untranslated region with four potential polyadenylation signal sites.

In order to determine the 3′-end of the cPR mRNA, an

```
-354  ACTGGAAAATGGAAGATGTATTCATGCTAACAGGGCGTACAGGAACTTAAACCTGAAAAACACAGCAAACGTCTGAAAGAGAGGGTGAAGAACCAGGAAGGAGAGGACATATCTTCCTCC

-234  CCGCGCTCCGTAACAGCGATTTGCCCAGCCTGAGTGCAGCCAGTATTGCACTCCTCGACCAGGACGGACACCTGCAAGAGCCGGGAGACGGGGTGAGATGGGGCTGCGAAGGCAGGAGGC

-114  GGCCGGGCGGCTAAGAGCGGAGCCGAGAGGGAACGGGAGCGGTGGGAGGGTCGGAGCAGGCGGAGCGAAGGAGGGGCCAGGCCGTGCCGGGAGCTGCCGGAGGGAGTCGCCGGTGCCCGA
                          TCGGCTCTCCCTTGCCCT oligo 1                                                                    +1
               BglI
   7  GAGCCGAGCGCCGAAGGGGCGAGAAGGCGGCTGCCTGCCGGGGGCAGCGGGGAAGCGGAGGGGCGAACCCGGGCGCCTCTCAGCAGTACGGGCGGCGGCGGACGGGGAGCGTCCAGGG
                                                                                    GAGAGTCGTCATGCCCGCCG oligo 2

 127  GAGCGCGGCGAGGAGCAGGGCGGGCAGCGGGGACGAGGCGCCGTCCCCGCCCGCCCTCTCGCTGCCCCTCGGAAGACGGCTGACGCTCAGATCGCTCTCCCCTCCGCAGGACCGTCCGG
                                                                                    ***         AGCGAGAGGGGAGGCGT oligo 3

 247  GCGGCGATCGGCGCACGGTCGGGGCTCCTCGGAACGCAGCGCCATGTGAACGGGAGCCCCGTCCGGGCCGCGAGGAGCCGCGCCGGAGGTTTGGCGGCGTCCAGCGCTGCTCCGGGCAGC

 367  ATG ACC GAG GTG AAG AGC AAG GAG ACC CGA GCG CCT TCG TCC GCC CGA GAC GGG GCG GTC CTG CTG CAG GCT CCC CCC AGC CGC GGG GAA
      Met Thr Glu Val Lys Ser Lys Glu Thr Arg Ala Pro Ser Ser Ala Arg Asp Gly Ala Val Leu Leu Gln Ala Pro Pro Ser Arg Gly Glu   30

 457  GCG GAG GGC ATC GAC GTC GCC CTG GAC GGG CTG CTC TAC CCC AGA AGC AGC GAC GAG GAG GAG GAG GAG GAG GAG AAC GAG GAG GAG GAG
      Ala Glu Gly Ile Asp Val Ala Leu Asp Gly Leu Leu Tyr Pro Arg Ser Ser Asp Glu Glu Glu Glu Glu Glu Glu Asn Glu Glu Glu Glu   60

 547  GAG GAG GAA GAA CCG CAG CAG CGG GAG GAG GAG GAA GAG GAG GAG GAG GAG GAT CGG GAC TGC CCC TCG TAC CGG CCG GGC GGC GGC TCG
      Glu Glu Glu Glu Pro Gln Gln Arg Glu Glu Glu Glu Glu Glu Glu Glu Glu Asp Arg Asp Cys Pro Ser Tyr Arg Pro Gly Gly Gly Ser   90

 637  CTC TCC AAG GAC TGT CTG GAC AGC GTC CTG GAC ACC TTC CTG GCG CCC GCA GCC CAT GCC GCG CCC TGG TCC CTA TTC GGG CCG GAG GTG
      Leu Ser Lys Asp Cys Leu Asp Ser Val Leu Asp Thr Phe Leu Ala Pro Ala Ala His Ala Ala Pro Trp Ser Leu Phe Gly Pro Glu Val  120

 727  CCG GAA GTG CCC GTC GCC CCG ATG AGC CGC GGC CCC GAG CAG AAG GCT GTG GAC GCC GGC CCG GGG GCT CCC GGT CCC TCG CAG CCG CGA
      Pro Glu Val Pro Val Ala Pro Met Ser Arg Gly Pro Glu Gln Lys Ala Val Asp Ala Gly Pro Gly Ala Pro Gly Pro Ser Gln Pro Arg  150

 817  CCC GGG GCT CCG CTG TGG CCG GGC GCC GAC TCC CTG AAC GTC GCC GTG AAA GCG CGT CCG GGA CCC GAG GAC GCC AGC GAG AAC AGA GCG
      Pro Gly Ala Pro Leu Trp Pro Gly Ala Asp Ser Leu Asn Val Ala Val Lys Ala Arg Pro Gly Pro Glu Asp Ala Ser Glu Asn Arg Ala  180

 907  CCG GGG CTG CCC GGC GCC GAG GAG CGG GGC TTC CCG GAG CGG GAC GCG GGG CCC GGA GAA GGC GGA CTG GCG CCC GCC GCG GCT GCT TCC
      Pro Gly Leu Pro Gly Ala Glu Glu Arg Gly Phe Pro Glu Arg Asp Ala Gly Pro Gly Glu Gly Gly Leu Ala Pro Ala Ala Ala Ala Ser  210

 997  CCG GCG GCC GTG GAG CCG GGC GCG GGG CAG GAC TAC CTG CAC GTG CCG ATC CTG CCG CTC AAC TCT GCC TTC CTG GCC TCG CGC ACG CGG
      Pro Ala Ala Val Glu Pro Gly Ala Gly Gln Asp Tyr Leu His Val Pro Ile Leu Pro Leu Asn Ser Ala Phe Leu Ala Ser Arg Thr Arg  240

1087  CAG CTG CTG GAC GTG GAG GCG GCG TAC GAC GGC AGC GCC TTC GGG CCG CGC TCC TCG CCT TCC GTC CCC GCC GCG GAC TTG GCG GAG TAC
      Gln Leu Leu Asp Val Glu Ala Ala Tyr Asp Gly Ser Ala Phe Gly Pro Arg Ser Ser Pro Ser Val Pro Ala Ala Asp Leu Ala Glu Tyr  270

1177  GGC TAC CCG CCG CCC GAC GGC AAG GAG GGC CCC TTC GCT TAC GGC GAG TTC CAG AGC GCG CTG AAG ATC AAG GAG GAG GGC GTC GGT CTG
      Gly Tyr Pro Pro Pro Asp Gly Lys Glu Gly Pro Phe Ala Tyr Gly Glu Phe Gln Ser Ala Leu Lys Ile Lys Glu Glu Gly Val Gly Leu  300

1267  CCC GCC GCT CCC CCC CCG TTC CTG GGC GCC AAG GCC GCC CCC GCC GAC TTC GCG CAG CCC CCG CGG GCC GGG CAG GAG CCC TCG CTG GAG
      Pro Ala Ala Pro Pro Pro Phe Leu Gly Ala Lys Ala Ala Pro Ala Asp Phe Ala Gln Pro Pro Arg Ala Gly Gln Glu Pro Ser Leu Glu  330

1357  TGC GTC CTC TAC AAG GCC GAG CCG CCC CTC CTG CCC GGC GCA TAC GGG CCG CCC GCG GCC CCC GAC AGC CTG CCG TCC ACC TCG GCC GCG
      Cys Val Leu Tyr Lys Ala Glu Pro Pro Leu Leu Pro Gly Ala Tyr Gly Pro Pro Ala Ala Pro Asp Ser Leu Pro Ser Thr Ser Ala Ala  360

1447  CCG CCC GGC CTC TAC TCG CCG CTG GGC CTT AAC GGG CAC CAC CAG GCG CTG GGC TTC CCC GCG GCC GTG CTC AAG GAG GGC CTG CCC CAG
      Pro Pro Gly Leu Tyr Ser Pro Leu Gly Leu Asn Gly His His Gln Ala Leu Gly Phe Pro Ala Ala Val Leu Lys Glu Gly Leu Pro Gln  390

1537  CTG TGC CCG CCC TAC CTC GGC TAC GTG CGG CCA GAC ACA GAA ACC AGC CAG AGC TCC CAG TAC AGC TTT GAA TCG CTA CCC CAG AAG ATT
      Leu Cys Pro Pro Tyr Leu Gly Tyr Val Arg Pro Asp Thr Glu Thr Ser Gln Ser Ser Gln Tyr Ser Phe Glu Ser Leu Pro Gln Lys Ile  420

1627  TGT CTC ATC TGT GGT GAT GAG GCT TCT GGT TGC CAC TAC GGA GTA CTC ACC TGT GGA AGC TGT AAA GTC TTC TTT AAA AGG GCA ATG GAA
      Cys Leu Ile Cys Gly Asp Glu Ala Ser Gly Cys His Tyr Gly Val Leu Thr Cys Gly Ser Cys Lys Val Phe Phe Lys Arg Ala Met Glu  450

1717  GGG CAG CAC AAC TAT TTA TGT GCT GGA AGA AAT GAC TGC ATA GTT GAT AAA ATT CGT AGG AAG AAC TGT CCA GCG TGT CGC TTG AGG AAG
      Gly Gln His Asn Tyr Leu Cys Ala Gly Arg Asn Asp Cys Ile Val Asp Lys Ile Arg Arg Lys Asn Cys Pro Ala Cys Arg Leu Arg Lys  480

1807  TGC TGT CAA GCT GGA ATG GTC CTG GGA GGT CGA AAA TTT AAA AAG CTT AAC AAA ATG AAG GTT GTG AGG ACA TTA GAT GTT GCA CTT CAG
      Cys Cys Gln Ala Gly Met Val Leu Gly Gly Arg Lys Phe Lys Lys Leu Asn Lys Met Lys Val Val Arg Thr Leu Asp Val Ala Leu Gln  510

1897  CAG CCA GCC GTG CTT CAG GAT GAA ACC CAA TCT CTA ACG CAA AGG CTG TCC TTT TCT CCA AAT CAA GAA ATA CCG TTT GTT CCC CCA ATG
      Gln Pro Ala Val Leu Gln Asp Glu Thr Gln Ser Leu Thr Gln Arg Leu Ser Phe Ser Pro Asn Gln Glu Ile Pro Phe Val Pro Pro Met  540

1987  ATA AGT GTT TTG CGA GGC ATT GAG CCC GAA GTT GTC TAT GCT GGT TAT GAC AAT ACA AAA CCT GAA ACA CCA AGT TCC TTG CTG ACC AGT
      Ile Ser Val Leu Arg Gly Ile Glu Pro Glu Val Val Tyr Ala Gly Tyr Asp Asn Thr Lys Pro Glu Thr Pro Ser Ser Leu Leu Thr Ser  570

2077  CTA AAT CAT CTT TGT GAG AGG CAA CTG CTT TGT GTA GTC AAG TGG TCA AAA TTG CTA CCA GGA TTT CGG AAT TTA CAT ATT GAT GAC CAG
      Leu Asn His Leu Cys Glu Arg Gln Leu Leu Cys Val Val Lys Trp Ser Lys Leu Leu Pro Gly Phe Arg Asn Leu His Ile Asp Asp Gln  600

2167  ATA ACT CTC ATC CAG TAT TCT TGG ATG AGT CTA ATG GTT TTT GCA ATG GGA TGG AGA TCA TAC AAA CAT GTC AGT GGT CAG ATG TTG TAT
      Ile Thr Leu Ile Gln Tyr Ser Trp Met Ser Leu Met Val Phe Ala Met Gly Trp Arg Ser Tyr Lys His Val Ser Gly Gln Met Leu Tyr  630

2257  TTT GCA CCG GAT CTG ATT CTG AAT GAA CAG AGG ATG AAA GAA TCA TCG TTC TAT TCA CTG TGT CTC TCC ATG TGG CAG CTC CCA CAG GAG
      Phe Ala Pro Asp Leu Ile Leu Asn Glu Gln Arg Met Lys Glu Ser Ser Phe Tyr Ser Leu Cys Leu Ser Met Trp Gln Leu Pro Gln Glu  660

2347  TTT GTC AGA CTC CAA GTT AGT CAA GAG GAG TTT CTA TGT ATG AAA GCA CTA CTG CTT CTC AAT ACA ATT CCT TTG GAA GGT CTA AGA AGT
      Phe Val Arg Leu Gln Val Ser Gln Glu Glu Phe Leu Cys Met Lys Ala Leu Leu Leu Leu Asn Thr Ile Pro Leu Glu Gly Leu Arg Ser  690

2437  CAA AGC CAA TTT GAT GAG ATG AGA ACG AGT TAC ATT AGA GAA CTG GTG AAG GCC ATT GGT TTG CGC CAG AAA GGT GTG GTG GCT AAC TCC
      Gln Ser Gln Phe Asp Glu Met Arg Thr Ser Tyr Ile Arg Glu Leu Val Lys Ala Ile Gly Leu Arg Gln Lys Gly Val Val Ala Asn Ser  720

2527  CAG CGC TTC TAT CAG CTT ACA AAA CTG ATG GAT TCC ATG CAT GAT CTA GTG AAA CAG CTC CAT CTC TTC TGT CTG AAC ACA TTT CTC CAG
      Gln Arg Phe Tyr Gln Leu Thr Lys Leu Met Asp Ser Met His Asp Leu Val Lys Gln Leu His Leu Phe Cys Leu Asn Thr Phe Leu Gln  750

2617  TCC CGT GCA CTG AGC GTT GAA TTC CCT GAG ATG ATG TCA GAG GTG ATT GCT GCA CAG CTG CCC AAG ATT CTG GCA GGG ATG GTG AAA CCT
      Ser Arg Ala Leu Ser Val Glu Phe Pro Glu Met Met Ser Glu Val Ile Ala Ala Gln Leu Pro Lys Ile Leu Ala Gly Met Val Lys Pro  780

2707  CTT CTC TTT CAC AAG AAG TGA AAT GTC TTT TCT CTT ATC GTA GAG ATG ATT TCT GGG TCG GTT TTT GTT TGT TTA TTT TGG TCA AGA TTT
      Leu Leu Phe His Lys Lys ***                                                                                              786

2797  TGCAATGTCAGGACTTCAGTGTATTTGTCATACCCTGCCTACTGCTTTATACTTGTAACATACACAAGCCATTTAAGCTTGATGTACATATTGTGAGTTCCTGATTCCTTAACTGCAAAA
2917  ATCACAACAGTCGTAAGAAGTGTTTTGTAAACTTGGGTAGATTGCTTTTTAGATATGCGTAAGAATGGCCGAATGAATAGAGTTTTCAGATTTCTAAGAACAGTTATCTTATACATTTTT
3037  CTTACATATTACACACTTTCAAATTTCTGATATGACATGGTGTAAACCTTTTAATACACTGTGTGCATGGGTACACGTGGGCATGTGCATGGTCCTAAACATTTTGAGGGAATTAAAATG
3157  TATATTGCCCACACTTTCTCACGGATCTTTGACTTTTTATTTACTAATCTCCTTCACAGAAATATCCTTTCAAAGTCAGCTGTTAAGTTTAAGTAGATATCACTGCAGTACCTTCATGGG
3277  TAACTGAACTGCTCTCTGCAAAATATGTAAATATGGATTCATGTTCTTCTCTGTGCTCATGCTGTATTGAAGTTCTGATATCAGTGTGCTTCTTTCTGCATCATTCAGTGCAGTTCTCCT
3397  TCAAAAATCAGAAAGCAGAAGAAGATTTCAGTATATGTTATACCAGAAGTAGGATGTATGTTAACCCATACAGCTTCAGTTTGTGTTATTTTCCTTCAAACTTTGGCACTTAGTAAAGCT
3517  AAAAAACATTGTTGTAATCTTAAATTTATCAACTTAAAAGTTATGTATTTCAGAAGTTTTAAATAGAAATGTTTTTCCTGACTACCACTCTCAATAATCTTGATGGTTTGAATTCCCGTG
3637  CTGAATACCCAACCAGCCACTACCATCAAATTTTAGAGAGGTTAGAAACAAAAAGAAGTGGAAAATATGACCACAAACAGTTTGATAAACATAACATTCATCTTAATACCATTTAGTGAA
3757  GCTACATATTTTTTTACTTTAAAGGCATGCTGTTATACTTGGAGTAACTGTATGAGGAGTTGGGCTAAACAAGCATTGTACATTCAAGAAGAATGTCATGCTTTTCATGGAAATGAAATGT
3877  TTGACTAAAAAATGATAGTGTGATGGTTGTTGTTCCTTTGGTATGGACATGAAGCTACTCATCTTTGAGGGAGTAAACTGCTGATATATTCAGACAAACAGGAATGGATAAAGCCCAAGA
3997  AGAATTTTGGATTATTAAAATCTGATAAATTTTTACCAGCTCACTAGAAATGATACCCAGCTTCTCTAGCACTAATCAACGTAGGAAACTGCAAGTCTTGAATAACTGCTGACCATACGC
4117  TATTGCTTTTCCCCATTTAGACACTGGATGCATCCATTGTGGTAGCATTCTACTTATGGTTTGTAACATGGCATGTATGGACTTGTCTTTGTCAGGTAAGCTAAGGCCCATCAAGAGGAAG
4237  CTGTGTAACAGTTAGGAATAAAAGCATCCTACATAGTACAAAACCCACTGGAGTTGTGGCAAAGAACTGACCCAAGGTAAAATGCTCTAAGCAGACTTTTGTGTTTGTTCACTAGGTTCA
4357  TAACTTTTAACTGCCCCAAAACAAGCTATTTCTTACTTATTCAATACAGTCTGACTTTCATCTGTGATCCTTAAAGCAATTGTGTTTTAATAAATAACACAAGATTCCGTTTCTAC-(A)21
```

oligo(dT)-primed λgt10 cDNA bank was constructed from size selected (>3.5 kb) hen oviduct poly(A)$^+$ RNA. Using a genomic probe encoding the 3′ half of the last exon, 22 clones were isolated, of which 12 were polyadenylated. All of those contained identical 3′ ends located downstream from the last polyadenylation signal site, strongly suggesting that the cPR mRNA ends at position 4472 (see Figure 1) and that the last polyadenylation signal site is by far the preferred one.

The region containing the 5′ end of the cPR mRNA was first localized by Northern blotting of poly(A)$^+$ RNA from laying hen oviduct using the three single-stranded probes D, E and F described in Materials and methods and in Figure 2C. Two clear bands of ~4.5 and 3.3 kb and a fainter band of 6.5 kb were revealed using probes D and E, whereas no RNA hybridizing with probe F could be detected (Figure 2B, lanes 2, 4 and 6 respectively). A mixture of HeLa cell total RNA and *Hin*dIII−*Pst*I genomic DNA fragment (Control, see Figure 2C) was run in parallel in lanes 1, 3 and 5. As expected only the control DNA fragment was revealed in these lanes. These results indicated that the 5′-end of the cPR mRNA was located within a few hundred bases upstream of the *Bgl*I(+20) site (Figure 2C).

The primer extension technique was then used to map more precisely the position of the cap site. A single-stranded ^{32}P-labelled primer cut at the *Bgl*I(+20) site (Primer, in Figure 2C) was synthesized by elongation of the synthetic oligonucleotide 2 (Figure 1), using the *Hin*dIII−*Pst*I genomic DNA fragment (Figure 2C) inserted in M13 as a template. When elongated in the presence of oviduct poly(A)$^+$ RNA with reverse transcriptase and cold dNTPs, this labelled primer yielded a major product (Figure 2A, lane 2) which mapped at the position of a C in the mRNA coding strand of the DNA [see the sequence ladder of the DNA obtained using the dideoxy method with [^{35}S]-dATP and the oligonucleotide 2 (Figure 1) as a primer]. Some additional bands were visible both upstream and downstream of the main band suggesting the existence of multiple initiation sites. The 5′-end of the mRNA was also mapped with S1 nuclease using a ^{32}P-labelled probe C (Figure 2C, and Materials and methods). After hybridization with oviduct poly(A)$^+$ mRNA and S1 nuclease digestion, a major S1 nuclease-resistant DNA fragment (Figure 2A, lane 4) migrated at the same position as the reverse transcriptase extended primer (compare lanes 2 and 4). In this case also, several minor bands, possibly corresponding to multiple initiation sites, were visible on the original autoradiogram. Under identical conditions, no primer extension nor S1 nuclease-resistant products were seen using HeLa cell poly(A)$^+$ RNA (lanes 1 and 3). Thus the 5′-end of the cPR mRNA is most probably localized in a region centered on the G indicated as +1 in Figures 1 and 2A and there may be multiple initiation sites. That the 5′-end of the cPR mRNA is located in this region was also supported by S1 nuclease mapping experiments (not shown) carried out with ^{32}P-labelled probes A and B (Figure 2C), which resulted in protected fragments of the expected lengths.

All of the above results lead us to conclude that the length of cPR mRNA is very close to 4.5 kb, which is in good agreement with the estimated size of the major poly(A)$^+$ RNA species which was revealed on Northern blots with cPR cDNA probes (see above). Results which will be published elsewhere indicate that the 3.3 kb poly(A)$^+$ mRNA corresponds to a truncated form of the cPR mRNA which has a 5′-end similar to that of the 4.5 kb species but is interrupted in the coding region after amino acid 451. The origin of the 6.5 kb species is unknown but it may correspond to a splicing intermediate. It is important to note that, with the exception of the cDNA clones which correspond to the 3.3 kb mRNA species, we did not find any cDNA clones with a sequence not co-linear with that of the 4.5 kb mRNA in any of the chicken oviduct cDNA libraries which have been used. In addition the study of genomic clones did not provide any evidence that the chicken genome may contain different PR genes (to be published elsewhere).

In vivo and in vitro expression of the cloned cDNA shows that it encodes the progesterone receptor form B

Two constructs, both containing the entire open reading frame, were made using the expression vector pKCR2 (Breathnach and Harris, 1983) and the cPR cDNA. cPR0 contains the cDNA from position +29 to 2921, and thus most of the sequences coding for the 5′-untranslated region of the mRNA, whereas this region has been deleted in cPR1 where the ATG initiating the open reading frame is preceded by an engineered Kozak consensus sequence (Kozak, 1984a; see also Materials and methods and Figures 1 and 3). Specific binding of the synthetic progestin R5020 was readily detected in the cytosolic fraction of HeLa cells or Cos-1 cells transfected with either cPR0 or cPR1 and grown in a steroid hormone-free medium (Materials and methods; data not shown). Depending on the transfection efficiencies, specific bindings of ~500 and 2000 fmol/mg protein were found in HeLa and Cos-1 cells, respectively. When the affinity of the 'cloned' receptor for progesterone was determined by Scatchard analysis (data not shown), a K_D value of 4.7 nM was found, similar to that of the 'natural' progesterone receptor [5 nM, (Schrader *et al.*, 1980)]. The 'tight' nuclear binding of the 'cloned' receptor in the presence of hormone was analysed by adding labelled R5020 to the medium of HeLa cells transfected with either cPR0 or cPR1. Under these conditions most of the progesterone receptor was tightly bound to the nuclei of the transfected cells, since at least 70% of the total specific R5020 binding could be extracted from these nuclei with 0.5 M, but not 0.15 M NaCl (data not shown).

The size of the protein encoded in cPR0 and cPR1 was determined by Western blotting analysis of cytosolic extracts from Cos-1 cells transfected with either cPR0 or cPR1. In both cases (Figure 4A, lanes 4 and 8) polyclonal antibodies directed against the hen oviduct cPR form B (Tuohimaa *et al.*, 1984), revealed a major band migrating at the same level as the 'natural' chicken PR form B with an apparent mol. wt of 109 kd (Figure 4A, lane

Fig. 1. The chicken progesterone receptor: full-length cDNA nucleotide sequence and deduced amino-acid sequence. The 5′-flanking sequence of the chicken PR gene is shown from −354 to the cap site (+1, arrow). The PR mRNA sequence begins at nucleotide +1 and ends at nucleotide +4472. The sequence was determined from two genomic clones spanning position −354 to +1088 and +3626 to a position ~1 kb downstream of +4472, from four overlapping cDNA clones corresponding to positions +29 to +1983, +1651 to +2635, +2563 to +2921, +2635 to the poly(A) tail and subclones derived therefrom. The numbers on the left refer to the position of the nucleotides and those on the right to that of the amino-acids. The three oligonucleotides (oligo 1, 2 and 3) used for probe synthesis (Materials and methods) are given below the corresponding sequence, the most 5′-end of the cloned cDNAs (close to the *Bgl*I site, underlined) is indicated by a filled triangle at +29. Asterisks indicate the in-frame termination codon upstream and downstream of methionine 1. The two only methionines present in the 5′-half of the open reading frame are boxed. An ATGTGA motif, present in the 5′-non-translated region is marked by dashes, a polyglutamic acid region is indicated by dots. Peptide sequences which were found in tryptic digests of homogeneous cPR form B (Simpson *et al.*, 1987) are underlined. The central cysteine-rich DNA binding domain is boxed. In the 3′-untranslated region four potential polyadenylation signal sites are underlined.

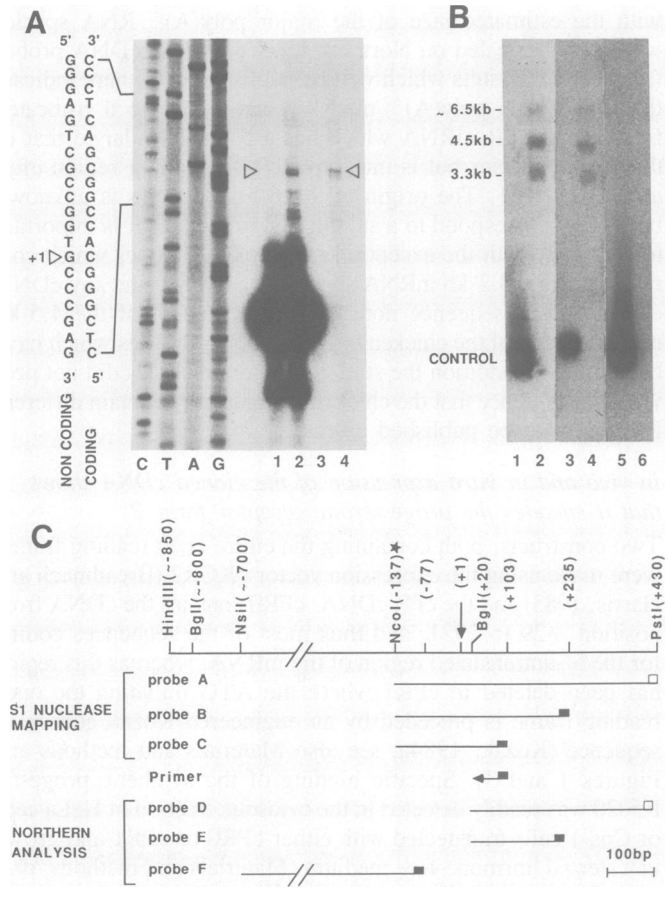

Fig. 2. Mapping of the cPR mRNA 5′ end. **A**. A 6% sequencing gel showing S1 nuclease (**lanes 3** and **4**) and primer extension (**lanes 1** and **2**) mapping carried out as described in Materials and methods using poly(A)$^+$ RNA from laying hen oviduct (**lanes 2** and **4**) or from HeLa cells (**lanes 1** and **3**). The **lanes CTAG** show the sequence ladder obtained using the dideoxy method and oligo 2 as primer elongated in the presence of [α^{35}S]-dATP. Open triangles indicate the position of the major transcription start site. **B**. Northern blot analysis performed according to Materials and methods using 15 μg of RNA per slot. **Lanes 1, 3** and **5** consist of total HeLa cell RNA mixed with 2 ng of the *Hin*dIII−*Pst*I fragment (see panel C) as a positive control; **lanes 2, 4** and **6** are poly(A)$^+$ RNA from oviduct. The probes used for hybridization were D (**lanes 1, 2**), E (**lanes 3, 4**) and F (**lanes 5, 6**). **C**. Map of the 5′-region of the cPR gene showing the probes and the Primer used in **A** and **B**. The M13 universal primer (open squares) was used to synthesize probes A and D, oligonucleotide 3 was used for the synthesis of probes B and E, oligonucleotide 2 for the synthesis of probe C and the Primer, and oligonucleotide 1 for probe F. The sequences of oligonucleotides (1 to 3) indicated by the filled squares are depicted in Figure 1. In all cases probe synthesis was carried out in the presence of [α-^{32}P]dATP, followed by digestion with *Nco*I (probes A, B and C), *Bgl*I (probes D, E and the Primer) and *Nsi*I (probe F) and purification as described in Materials and methods. The *Nco*I site (marked with an asterisk) at −137 was introduced by site-directed mutagenesis.

3). Note that only trace amounts of protein are detected in lanes 4 and 8 with a migration similar to that of the 'natural' cPR form A (79 kd, see lane 1). That the 'cloned' protein migrating with an apparent mol. wt of 109 kd was the cPR receptor form B, was further supported by electrophoretical analysis of the R5020 photoaffinity-labelled proteins in the cytosol of Cos-1 cells transfected with cPR0. A single band (Figure 4b, lane 3) is present on the autoradiogram, migrating at the same position as that of the photoaffinity-labelled partially purified 'natural' cPR form B (Figure 4B, lane 2). Again, no obvious band corresponding to the 'natural' cPR form A (lane 1) could be detected in the cytosol of the transfected Cos-1 cells.

The protein encoded in the cPR cDNA was also characterized using an *in vitro* transcription/translation system. The cPR cDNA sequence from +314 to +2921 was inserted in the pGEM vector to yield cPR0G (Materials and methods), where it could be transcribed *in vitro* under the control of the T7 promoter. cPR1G which contains the cPR sequence present in cPR1 was constructed similarly. The *in vitro* synthesized cPR mRNA was then translated in a rabbit reticulocyte lysate in the presence of [^{35}S]-methionine. Identical results were obtained using either cPR0G or cPR1G (not shown). The translation of cPR0G RNA resulted in the appearance of several major proteins (bands 1−4 in Figure 5, compare lanes 1 and 2). The largest protein had an apparent mol. wt of 109 kd and migrated at the same level as the R5020 photoaffinity-labelled 'natural' cPR form B (not shown in this figure). Note that, in contrast to the results obtained *in vivo* with cPR0 (see above), a protein migrating with an apparent mol. wt of 79 kd (band 2), at the same level as the R5020 photoaffinity-labelled 'natural' cPR form A (not shown in this figure), could be seen amongst the *in vitro* transcription products (see below).

Functional analysis of the chicken progesterone receptor

An amino acid sequence comparison between the chick progesterone receptor and its human (Misrahi *et al.*, 1987 and A.Krust, U.Stropp, P.Kaster and P.Chambon, unpublished results) and rabbit (Loosfelt *et al.*, 1986) counterparts reveals two highly conserved regions. The cysteine-rich region C spanning from amino acid number 410 to 495 (see Figure 1 and top of Figure 3) exhibits a 100% conservation, whereas region E from amino acid 540 to the C-terminus is 87% conserved, when compared with the PR of mammals whose ancestors diverged from those of birds ~100 million years ago. These two regions are separated by a region D of lower homology (~60%). On the other hand, the N-terminal region [termed A/B by analogy with the corresponding region of the oestrogen receptor (Krust *et al.*, 1986)] of these PRs diverges very significantly both in sequence and length (556, 556 and 409 amino acids for the human, rabbit and chicken PRs respectively). It is in fact this region which is responsible for the size difference between the cPR and mammalian PRs. As reviewed and discussed elsewhere, the highly conserved regions C and E of the progesterone receptor can be easily aligned with regions similarly located in the receptor of other members of the steroid/thyroid hormone receptor supergene family, whereas no conservation could be detected in regions A/B and D (Krust *et al.*, 1986; Green *et al.*, 1987; Gronemeyer *et al.*, 1987). Functional studies of the oestrogen (Kumar *et al.*, 1986) and glucucorticoid (Giguère *et al.*, 1986; Danielson *et al.*, 1986; Godowski *et al.*, 1987; Hollenberg *et al.*, 1987; Miesfeld *et al.*, 1987; Rusconi and Yamamoto, 1987) receptors have shown that the 66 amino acid-long core of region C, which is characterized by a specific cysteine-repeat motif (amino acids 421 to 486 in the cPR, see the boxed sequence in Figure 1 and also top of Figure 3) contains the domain responsible for 'tight' nuclear binding and specific recognition of the hormone responsive elements of the target genes (Green and Chambon, 1987; Kumar *et al.*, 1987 and V.Kumar, S.Green, G.Stack, M.Berry, J.R.Jin and P.Chambon, submitted). It has also been shown that region E of these receptors contains the hormone-binding domain and that this domain can function independently (Kumar *et al.*, 1986; Green *et al.*, 1987; Gronemeyer *et al.*, 1987).

Based on these results, we constructed the expression vector cPR3 which contains the C, D and E regions (see Figure 3 and

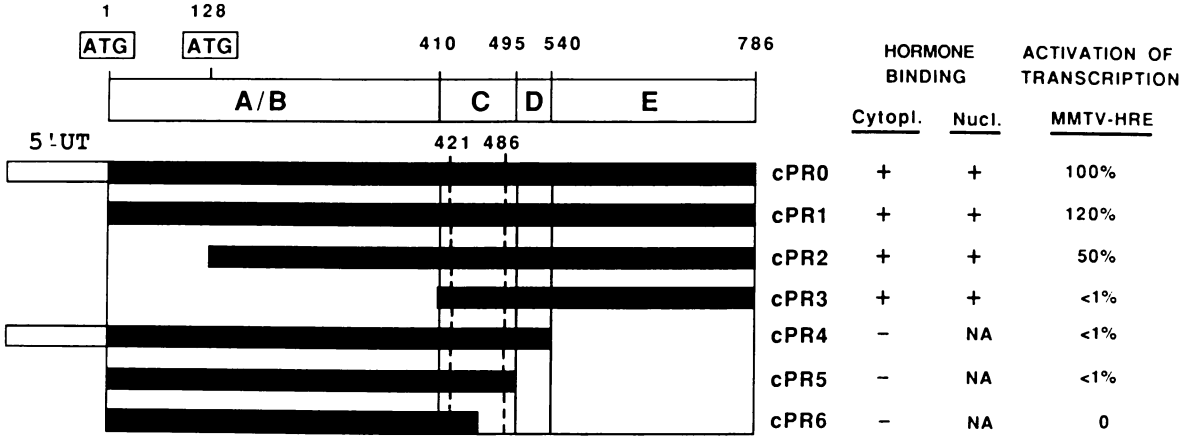

Fig. 3. Domain structure and functional properties of the cPR and its deletion mutants. Based on an amino acid sequence comparison with its human (Misrahi *et al.*, 1987) and rabbit (Loosfelt *et al.*, 1986) counterparts, the cPR has been subdivided into two highly conserved regions C (AA 410-495) and E (AA 540-786) exhibiting 100% and 87% homology with the mammalian PRs, respectively, separated by a region D of lower homology (~60%) and a N-terminal region A/B, differing considerably both in length and amino acid sequences from the mammalian progesterone receptors. The nomenclature of these regions is in analogy to Krust *et al.* (1986), the amino acid position of their boundaries is given at the top, as well as the two only ATGs present in region A/B. The central cysteine-rich core of region C (amino acids 421 to 486, see text) is indicated by dashed lines. A solid black bar represents the cPR sequences present in the deletion mutants (cPR2 to cPR6) with the deleted portion as a gap. The wild-type cPR sequence, cPR0, contains all 786 amino acids as well as the 5' untranslated sequence (5'-UT) downstream of position +29 (see Figure 1), indicated as an open box. The same 5'-region is also present in cPR4 whereas it is deleted in the other recombinants. The mutants cPR1 to cPR3 and cPR5 to cPR6 all contain 'consensus' Kozak sequences (Materials and methods) at their translational start sites. On the right side under HORMONE BINDING the ability of the various mutants to bind either progestins in cytoplasmic extracts (Cytopl.) or the ability of the progestin–receptor complex to bind tightly to HeLa cell nuclei (Nucl.) is shown (Materials and methods). (+) indicates wild-type PR characteristics, (−) indicates the absence of specific progestin binding. The ability of cPR wild type and mutant constructs to trans-activate transcription by interacting with the MMTV HRE is shown on the very right (ACTIVATION OF TRANSCRIPTION). The data which correspond to the average of at least three independent quantitative S1-nuclease analyses with different plasmid preparations (see Materials and methods and text), were obtained from densitometric scans of the autoradiograms and were corrected for transcription from the cotransfected internal control recombinant pG1B (see also Figure 6B). The values are relative to cPR0 taken as 100%. NA, not applicable.

Materials and methods). When transfected into HeLa cells grown in an hormone-free medium, cPR3 resulted in the appearance of a truncated cPR protein which bound progesterone with 'wild type' characteristics indistinguishable from those of the protein encoded in cPR0 or cPR1 (see above; Figure 3 'cytopl.' and data not shown). 'Tight' nuclear binding was also observed with the truncated receptor encoded in cPR3 (Figure 3 'nucl.' and data not shown). These results suggest very strongly that regions C and E of the cPR have the same functions as those previously established for the corresponding domains of the oestrogen and glucocorticoid receptors (see above). In particular it appears that the highly conserved region E can function as an independent progestin-binding domain because expression of the isolated region E resulted in the appearance of a truncated protein capable of binding progestins (data not shown). As expected no hormone binding could be detected in cells transfected with recombinants in which region E was deleted (cPR4, cPR5 or cPR6, see below and Figure 3).

It has been shown that not only the glucocorticoid receptor (GR), but also the PR can activate initiation of transcription *in vivo* from the MMTV LTR by interacting with its hormone-responsive element (HRE) (Cato *et al.*, 1986; 1987). This observation prompted us to investigate whether the 'cloned' cPR could trans-activate transcription from the MMTV LTR. cPR0 and cPR1 were transfected in HeLa cells grown in the presence or absence of R5020, together with the MMTV-CAT reporter gene (Cato *et al.*, 1987). A clear hormone-dependent stimulation of CAT expression was observed with cPR0, cPR1 or cPR2 (Figure 6A, compare lanes 3, 5 and 7 with lanes 4, 6 and 8; see also lanes 1 and 2 which correspond to transfections with the parent expression vector pKCR2). The results shown in Figure 6A indicate also that the truncated cPRs encoded in cPR3 (lanes 9 and 10) and cPR5 (lanes 11 and 12) had a drastically decreased stimulatory activity which, as expected, was independent of the presence of the hormone in the case of cPR5. No stimulation could be detected when cPR6 (lanes 13 and 14) was used. Similar results were obtained when transfections were performed in Cos-1, instead of HeLa cells (data not shown). These decreases in stimulatory activity were not due to reduced amounts of truncated receptor proteins in the transfected cells as shown by hormone binding in the case of cPR3 (data not shown) and Western blot analysis of the cytosolic proteins of Cos-1 cells transfected with either cPR4, 5 or 6 [Figure 4A, lanes 5, 6 and 7 respectively, in which truncated proteins with the expected size were present in amounts similar to those observed for the proteins encoded in cPR0 (lane 8), cPR1 (lanes 4) and cPR2 (lanes 2)].

In order to quantify more rigorously the transcriptional stimulatory activity of the various deletion mutants, RNA was determined by quantitative S1 nuclease analyses using an MMTV LTR-based reporter recombinant in which the CAT sequence was replaced by a promoterless rabbit β-globin gene (MMTV-globin or MG, see Materials and methods). In addition, in order to be able to correct for variations in transfection efficiencies, an internal control recombinant (pG1B) containing the rabbit β-globin gene and its promoter was co-transfected. Using this system, the RNAs initiated at the MMTV LTR startsite (+1 MG in Figure 6B) and at the globin startsite of pG1B (+1 pG1B) could be quantitatively determined by scanning the autoradiograms of polyacrylamide-urea gels, following hybridization with a single stranded ^{32}P-labelled DNA probe and nuclease S1 digestion (Materials and methods). The results of a representative transfection experiment in HeLa cells is displayed in Figure 6B, and the average values obtained by scanning at least three similar autoradiograms are given in Figure 3, relative to those obtained with cPR0 (taken as 100%, after correction for transcription from the internal control pG1B). A very strong stimulation (> 100-fold)

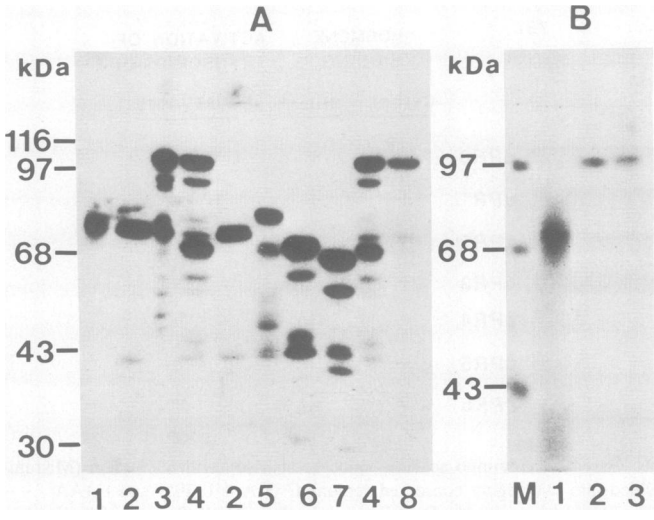

Fig. 4. Analysis by immunoblotting and photoaffinity labelling of the 'cloned' cPR and cPR deletion mutants. **A.** Immunoblot of partially purified (Gronemeyer *et al.*, 1985) hen oviduct progesterone receptor forms A (**lane 1**) and B (**lane 3**) and of crude cytosol prepared from Cos-1 cells transiently transfected with cPR0 (**lane 8**), cPR1 (**lanes 4**), cPR2 (**lanes 2**), cPR4 (**lane 5**), cPR5 (**lane 6**) and cPR6 (**lane 7**). Approximately 9 μg, 12 μg and 60 μg of total protein of the oviduct partially purified forms A and B, and of the Cos-1 cell cytosol, respectively, were separated on a 7.5% SDS–polyacrylamide gel and processed for immunoblotting as described (Jeltsch *et al.*, 1986) using polyclonal antiserum directed against the chick oviduct cPR form B (Tuohimaa *et al.*, 1984) at a dilution of 1:200. The mol. wts indicated on the left are derived from prestained mol. wt markers (BRL) transferred onto the same blot left and right of the sample lanes. **B.** Partially purified chick oviduct cPR form A (**lane 1**) and B (**lane 2**) and crude cytosol of Cos-1 cells transiently transfected with cPR0 (**lane 3**) were incubated with [^{3}H]R5020 and photoaffinity labelled by UV-irradiation as described (Gronemeyer *et al.*, 1985). The cross-linking products were separated on a 7.5% SDS–polyacrylamide gel and visualized by fluorography using EN^{3}HANCE(NEN). Exposure time was 3 weeks. The size of ^{14}C-labelled protein markers (M) is given on the left.

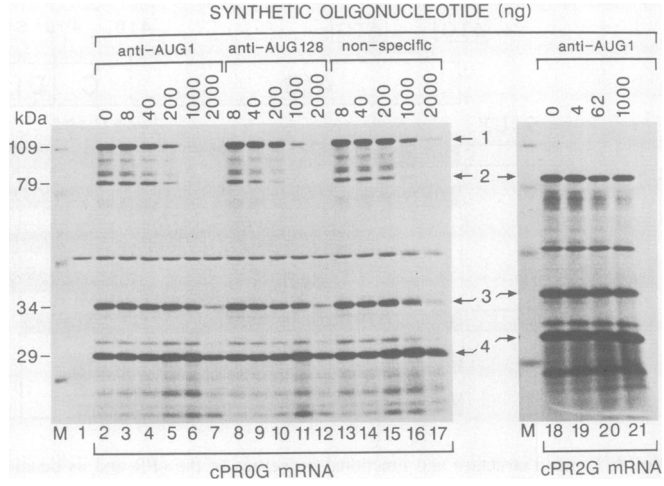

Fig. 5. *In vitro* translation of the chicken progesterone receptor and blocking of initiation at AUG 1 and AUG 128. The mRNA corresponding to the cPR sequences +314 to +2921 or +748 to +2921 (see Figure 1) was synthesized *in vitro* from cPR0G and cPR2G respectively, as described in Materials and methods. Identical aliquots of the same cPR0G-derived RNA preparation (~100 ng) were hybridized with no (**lane 2**) or 8 ng (**lanes 3, 8, 13**), 40 ng (**lanes 4, 9, 14**), 200 ng (**lanes 5, 10, 15**), 100 ng (**lanes 6, 11, 16**) or 2000 ng (**lanes 7, 12, 17**) of one of the three synthetic oligonucleotides (as indicated) corresponding to sequences surrounding the first (anti-AUG 1) or the second (anti-AUG 128) ATG present in the 5′-region of the cDNA (see Figure 1), or to an unrelated sequence (non-specific) before translation *in vitro* in the presence of [^{35}S]methionine (see Materials and methods). The labelled proteins were separated by SDS–gel electrophoresis and visualized by fluorography. **Lane 1** corresponds to an *in vitro* translation reaction where no RNA was added. The *in vitro* translation products obtained with cPR2G-derived mRNA, without (**lane 18**) or with previous hybridization with 4 ng (**lane 19**), 62 ng (**lane 20**) or 1000 ng (**lane 21**) of oligonucleotide anti-AUG 1 are also shown. The protein bands corresponding to translation inhibition at Met-1, Met-128, Met-486 and Met-540 are numbered 1, 2, 3 and 4, respectively. Their mol. wts are determined as 109 kd (band 1), 79 kd (band 2), 34 kd (band 3) and 29 kd (band 4), as indicated on the left side. The ^{14}C-labelled molecular markers [M] are the same as in Figure 4B.

was observed with cPR0 in the presence of the progestin R5020. In fact this stimulation was ~4-fold higher than that obtained under identical conditions (but in the presence of triamcinolone acetonide) with HG1, a pKCR2-based vector expressing the human glucocorticoid receptor (Kumar *et al.*, 1987).

The stimulatory activity of cPR1 was very close to that of cPR0, whereas that of cPR2 which encodes a truncated receptor which may correspond to the 'natural' cPR form A (see below) was decreased by approximately 2-fold. In agreement with the results of the CAT assay, deletion of the A/B region down to amino-acid 405 resulted in a decrease of stimulation of transcription by at least two orders of magnitude (cPR3 in Figure 6B, lanes 11 and 12 and Figure 3). This is particularly striking because the truncated PR protein encoded in cPR3 bound progesterone with 'wild' type characteristics and was 'tightly' bound to the nucleus in the presence of progestins (see above). Finally, very little (cPR4 and cPR5, Figure 6B, lanes 13–15 and Figure 3) or no (cPR6, Figure 3 and data not shown) stimulation of transcription was observed with the other deletion mutants, as expected from the results obtained with MMTV-CAT.

Possible origin of the chicken progesterone receptor form A

Previous work from our laboratory has provided strong evidence that the progesterone receptor form B (apparent mol. wt 109 kd) and form A (apparent mol. wt 79 kd) both of which can be found in chicken oviduct, are structurally and immunologically closely related (Gronemeyer *et al.*, 1983; Gronemeyer *et al.*, 1985 and see Introduction). The present cloning of the cPR cDNA as well as that of the corresponding genomic sequence (to be published elsewhere) exclude that forms A and B could be the products of different genes. Furthermore we found no evidence that form A could be encoded in a different mRNA generated by alternative termination of transcription or alternative splicing. In addition, since the 'natural' forms A and B bind progestin and the hormone-binding domain is located in the C-terminal region (see above), it is most probable that the difference between the two proteins resides in their N-terminal region. Therefore, two basic possibilities have to be considered: (i) form A corresponds to a protein which is initiated from an internal AUG; or (ii) form A results from proteolysis of form B in its N-terminal region.

There is only one AUG codon (AUG 128, boxed in Figure 1) between the AUG which initiates the open reading frame of the cPR form B (AUG 1) and the beginning of domain C. Initiation at AUG 128 would yield a protein with a calculated mol. wt of 72 kd, close to the apparent mol. wt of the 'natural' form A. The expression vector cPR2 (Figure 3 and Materials and methods) whose initiation codon corresponds to AUG 128, and its cPR2G counterpart for *in vitro* transcription/translation (Materials and methods), were constructed to study whether the 'natural' form A may correspond to a protein initiated at AUG 128. Expression both in Cos-1 cells *in vivo* (Figure 4A, lanes 2) and *in vitro* (Figure 5, lane 18) resulted in a protein with an

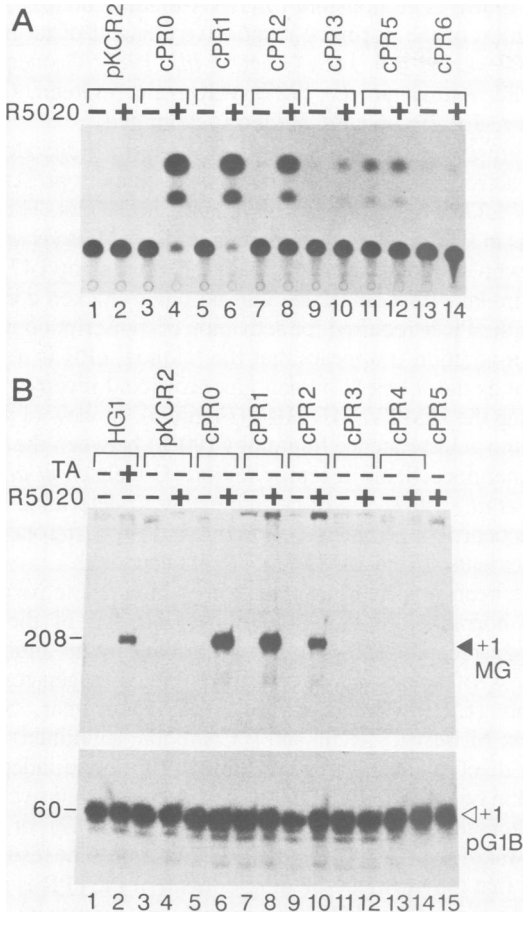

Fig. 6. Activation of transcription by the cPR and its deletion mutants. (**A**) CAT assay using the cytosol of HeLa cells transiently transfected with 1 μg of the cPR expression vectors (**lanes 3–14**, as indicated) or with the parent vector pKCR2 (**lanes 1** and **2**) and grown in the presence (+) or absence (−) of 5 nM R5020. The reporter plasmid used was MMTV-CAT (Cato *et al.*, 1987). The amount of extracts for the CAT assay were normalized by measuring the β-galactosidase activity generated from the co-transfected vector pCH110 (Pharmacia) which encodes the *E. coli* β-galactosidase. (**B**) Quantitative S1 nuclease analysis. The human glucocorticoid expression vector (HG1, **lanes 1** and **2**), the parent vector pKCR2 (**lanes 3** and **4**) or the cPR expression vectors as indicated (**lanes 5–15**) were co-transfected in HeLa cells together with a MMTV-globin reporter recombinant (MG) and the internal reference vector pG1B (for description of these recombinants, see Materials and methods). The cells were grown in the presence (+) or absence (−) of 5 nM progestin (R5020) or triamcinolone acetonide (TA) as indicated. After isolation of the RNA, hybridization and S1 nuclease treatment were resolved on denaturing polyacrylamide gels. The fragments seen on the autoradiogram correspond to RNA initiated at the startsite of transcription in the MMTV LTR (208 nucleotides in length, indicated by +1MG) and in the internal control recombinant (60 nucleotides in length indicated by +1 pG1B).

apparent mol. wt of 79 kd, identical to that of the 'natural' form A revealed either by immunoblotting (Figure 4A, lane 2) or photoaffinity labelling with R5020 (data not shown). On the other hand, a protein with a similar apparent mol. wt could be detected on immunoblots of Cos-1 cells transfected with cPR0 or cPR1 (see above and Figure 3) only after a long exposure of the autoradiograms [Figure 4A, lane 4 and 8; note that both forms A and B can be detected in equal amounts in laying hen oviduct cytosol using photoaffinity labelling (Gronemeyer *et al.*, 1983; Gronemeyer and Govindan, 1986)]. In contrast, a protein migrating similarly to the 'natural' form A was readily visible on the autoradiograms of the *in vitro* products of transcription/translation of cPR0G (Figure 5, lane 2) or cPR1G (data not shown). From these results we concluded that: (i) the protein initiated at AUG 128 has an apparent mol. wt identical to that of the 'natural' form A; (ii) such a protein is also present in the *in vitro* translation products of an mRNA coding for the form B, whereas it is barely detectable in the products synthesized from such an mRNA in Cos-1 cells *in vivo*.

To investigate whether the putative form A synthesized *in vitro* could be internally initiated at AUG 128, we studied the effect on its synthesis of hybridizing oligonucleotides complementary to the regions containing AUG 1 or AUG 128 (Figure 5). Densitometric scanning of this and similar autoradiograms indicated that the synthesis of the 79 kd form A from cPR0G mRNA was 50% inhibited by the oligonucleotide anti-AUG 128 (lane 10), under conditions where it was not specifically inhibited by either the oligonucleotide anti-AUG 1 (lane 5) or a control non-specific oligonucleotide (lane 15) (~25% inhibition in both cases; see also lane 20). Furthermore, the band 1 (form B)/band 2 (form A) ratio decreased in the presence of increasing amounts of the oligonucleotide anti-AUG 1, whereas the reverse situation was observed with the oligonucleotide anti-AUG 128 and no change of this ratio was observed with the control non-specific oligonucleotide. It appears therefore that at least under these *in vitro* conditions, internal initiation from AUG 128 can occur resulting in a protein with the same mol. wt as the 'natural' form A. Internal initiation at AUG 486 and AUG 540 would also account for the synthesis of the 34 kd and 29 kd *in vitro* translation products (bands 3 and 4 in Figure 5), which have no known *in vivo* counterparts (Gronemeyer and Govidan, 1986). On the other hand, it is most probable that the bands migrating between the 109 kd and 79 kd proteins correspond to proteolytic products of the form B, since in all cases their relative intensity followed closely that of the 109 kd form B.

There is a marked discrepancy between the mol. wt of the cPR form B (85 941 daltons) deduced from its amino acid sequence and its apparent mol. wt estimated from its electrophoretic migration (109 kd). On the other hand, there is a good agreement between the cDNA derived (72 kd) and the observed (79 kd) mol. wts of the protein encoded in cPR2 which is initiated at AUG 128. This observation indicates that the above discrepancy is most probably related to an amino acid sequence located upstream of methionine 128. This conclusion is further supported by the observation that the apparent mol. wts of the proteins encoded in cPR4 (Figure 4A, lane 5), cPR5 (Figure 4A, lane 6) and cPR6 (Figure 4, lane 7) are also larger by ~21 kd than expected from the mol. wts deduced from their amino acid sequence (81 kd, 73 kd and 69 kd versus 60 kd, 52 kd and 48 kd, respectively). The unusual migration of these proteins, as well as that of the form B, is most probably due to the presence of the polyglutamic acid stretch which is highlighted by dots in Figure 1 (amino acids 49 to 77). Similar abnormal electrophoretic behaviours, due to the polyglutamic acid sequences have indeed been observed for other proteins (Dingwall *et al.*, 1987; Earnshaw *et al.*, 1987).

Discussion

The cloned cDNA encodes the form B of the chicken progesterone receptor

Two 'natural' forms of the progesterone receptor with apparent mol. wts of 109 kd and 79 kd have been previously found in approximately equimolar amounts in the chicken oviduct (see Introduction). Our present results indicate that both of them are encoded in a single 4472 nucleotide-long mRNA whose *in vivo*

and *in vitro* translation yields a protein which is indistinguishable from the 'natural' form B by electrophoretic migration, binding of progestins and 'tight' nuclear association following binding of the hormone *in vivo*. The discrepancy between the present mol. wt of cPR form B deduced from its amino acid sequence (85 941 daltons) and its apparent mol. wt (109 kd) estimated from its electrophoretic migration in SDS−polyacrylamide gels (Schrader *et al.*, 1980; Gronemeyer and Govindan, 1986) appears to be due to the presence of a polyglutamic acid stretch in its N-terminal region. The possible significance of this sequence is unknown, but it is interesting to note that various amino acid homopolymers (usually polyglutamine) have been found in other proteins involved in development, such as the homeogene proteins (the so-called OPA sequences; for refs., see Wharton *et al.*, 1985; Laughon *et al.*, 1985) and in some nuclear proteins (Dingwall *et al.*, 1987; Earnshaw *et al.*, 1987). However this polyglutamic acid sequence is not present in the rabbit (Loosfelt *et al.*, 1986) nor in the human (Misrahi *et al.*, 1987; unpublished results from our laboratory) progesterone receptors.

In vivo expression in Cos-1 cells of a cloned cPR cDNA deleted for the sequences located upstream from the AUG coding for methionine 128 resulted in the appearance of a protein indistinguishable from the natural cPR form A by both its electrophoretic migration and hormone binding properties. From these results, it appears that the N-terminus of the 'natural' cPR form A is either at, or very close to, methionine 128. Thus the natural form A could correspond to either internal initiation of translation at the AUG coding for methionine 128 or to a specific proteolysis of form B occurring very close to this methionine. Our *in vitro* translation studies indicate that a protein with an electrophoretic mobility identical to that of the 'natural' cPR form A can result from translation of form B mRNA by initiation at the internal AUG coding for methionine 128. However *in vivo* expression in Cos-1 cells of the cloned cPR cDNA encoding the form B yielded only very low amounts of a protein migrating like the 'natural' cPR form A. Thus this internal initiation of translation must be a cell-specific process if it is responsible for the synthesis of the oviduct cPR form A as suggested by our *in vitro* results. The present study does not exclude that proteolysis could be responsible for the generation of the 'natural' form A; however our results indicate that this process, if occurring, must also exhibit cell-specificity. Determination of the N-terminal amino acid of the oviduct cPR form A is obviously required to definitely establish how this cPR form is generated from the same mRNA species as that coding for form B. We note in this respect that both the human and rabbit (Loosfelt *et al.*, 1986; Misrahi *et al.*, 1987) PR mRNAs contain an internal methionine equivalent to that present at amino acid position 128 in the cPR. Interestingly, although both A and B forms have been found 'naturally' in human breast cancer cells (Horwitz *et al.*, 1985), Milgrom's group has shown conclusively that in the rabbit uterus form A is an *in vitro* proteolytic product (Logeat *et al.*, 1985).

Functional domains of the chicken progesterone receptor

The striking amino acid conservation of domains C and E of the chicken, rabbit (Loosfelt *et al.*, 1986) and human (Misrahi *et al.*, 1987) progesterone receptors suggests that they have the same function as the homologous domains C and E of the oestrogen and glucocorticoid receptors (for refs. and reviews, see Gronemeyer *et al.*, 1987). Our results indicate that the cPR region E does indeed contain the progestin-binding domain and suggest that region C is responsible for 'tight' nuclear binding. It is therefore likely that the highly conserved region C, whose sequence has the potential to form DNA-binding fingers (Jeltsch *et al.*, 1986), is responsible for DNA-binding and the specific recognition of the hormone responsive element of target genes (for reviews and refs., see Green *et al.*, 1987; Gronemeyer *et al.*, 1987; Kumar *et al.*, 1987; V.Kumar, S.Green, G.Stack, M.Berry, J.R.Jin and P.Chambon, submitted).

Our present study shows that the cPR form B encoded in the cDNA activates the hormone responsive element present in the MMTV LTR when expressed in human HeLa cells. This result, which is in keeping with the previous studies of Cato *et al.* (1986, 1987) who showed that the human PR present in T47-D cells has a similar effect, indicates that no other chicken component besides the PR is required for activation of transcription in human cells. In addition it appears that the 'cloned' cPR is at least as efficient as the 'cloned' human glucocorticoid receptor (GR) at activating the MMTV LTR HRE (Figure 6). In view of the striking amino acid sequence homology (90%) between the PR, GR and mineralocorticoid receptor regions C (Jeltsch *et al.*, 1986; Arriza *et al.*, 1987) it will be interesting to determine whether the glucocorticoid, mineralocorticoid and progesterone responsive elements of the MMTV LTR are identical.

It has been reported that mutations which delete part, or all, of the hormone binding domain (regions E) of the human oestrogen receptor (hER) are accompanied by an almost complete loss of activation of two different oestrogen-responsive elements (ERE) (Kumar *et al.*, 1987; V.Kumar, S.Green, G.Stack, M.Berry, J.R.Jin and P.Chambon, submitted) whereas similar deletions result in a constitutive (hormone-independent) activation of the MMTV LTR glucocorticoid-responsive element by either the human (Hollenberg *et al.*, 1987) or the rat (Godowski *et al.*, 1987) 'cloned' GR. Our present results show clearly that the hormone-binding domain of the cPR is essential for activation of the MMTV progesterone responsive element (PRE). Deletions of the A/B region of the oestrogen and glucocorticoid receptors have led to various effects on activation of transcription. Deletions in this region in the human (Hollenberg *et al.*, 1987) GR resulted in up to 90% decrease of stimulation of transcription from the MMTV LTR. A similar decrease was observed with an hER deleted for region A/B when using the human pS2 gene ERE, whereas there was no decrease using a *Xenopus* vitellogenin gene ERE (Kumar *et al.*, 1987). It is clear from our present results that some sequences located in the cPR region A/B, downstream from methionine 128, are essential for activation of the MMTV PRE (see cPR2 and cPR3, in Figure 3). Since there is a significant conservation in the amino acid sequence of the chicken, human and rabbit PRs for ~40 amino acids upstream from the deletion end point of cPR3, it will be interesting to investigate whether this region is important for activation of transcription. Finally we note that the 'cloned' PR form A (cPR2 in Figures 3 and 6) is almost as efficient as the 'cloned' form B at activating transcription from the MMTV PRE, indicating that, if actually made in the chicken oviduct, the 'natural' cPR form A may activate transcription from target genes.

Materials and methods

Isolation of cDNA clones and sequencing

Overlapping clones covering the entire cPR mRNA were isolated from the λgt10 and λgt11 libraries described earlier (Jeltsch *et al.*, 1986). The isolation of clones corresponding to the cPR gene will be reported in detail elsewhere. For the determination of the mRNA 3′ end region an oligo(dT)-primed λgt10 library was constructed using size-selected hen oviduct mRNA (>3.5 kb). Screening of 2×10^5 recombinants with a genomic probe covering the 3′ half of the last exon gave 22 positive clones; 12 of these contained a poly(A) sequence and identical 3′ ends. The remaining 10 clones lacked parts of the 3′ untranslated region.

Sequencing of overlapping DNA fragments was performed on both strands by the Sanger dideoxy technique. GC-rich regions were sequenced according to Mizusawa *et al.*, (1986) and confirmed by the chemical degradation method (Maxam and Gilbert, 1980).

S1 nuclease and primer extension analyses

Probes for the determination of the mRNA 5'-end by S1 nuclease mapping and primer extension analysis were synthesized using a single-stranded M13mp19 clone containing the *Hin*dIII−*Pst*I gene fragment located in the 5'-end region of the gene (see Figure 2C). Second strand synthesis was primed with synthetic oligonucleotides (probe A: M13 universal primer; probe B: oligo 3 in Figure 1; probe C and the Primer: oligo 2 in Figure 1) in the presence of 1 μM [α-^{32}P]-dATP (specific activity 800 Ci/mmol) as a radioactive precursor and DNA polymerase I (Klenow fragment, 10 units, Appligene). After cleavage with *Nco*I (S1-nuclease probes) or *Bgl*I (Primer), the labelled DNA was isolated on a 4% denaturing polyacrylamide gel and electroeluted. Hybridizations were performed overnight at 50°C in 20 μl of 400 mM NaCl, 40 mM Pipes pH 6.5, 50% formamide with 10 μg of hen oviduct poly(A)$^+$ RNA and about 200 000 c.p.m. of probe (or Primer). S1 nuclease (300 units, Appligene) digestions were carried out in 200 μl of 30 mM sodium acetate pH 4.5, 300 mM NaCl and 3 mM $ZnCl_2$ for 2 h at 25°C. Primer extension was in 50 mM Tris−HCl pH 8.3, 8 mM $MgCl_2$, 250 μM of each dNTP using 50 units of reverse transcriptase (Molecular Genetic Resources) in a total volume of 200 μl.

mRNA analysis by Northern blotting

Total RNA was isolated according to Auffray and Rougeon (1980), purified by oligo(dT)-cellulose chromatography, electrophoresed on 1.3% agarose-formaldehyde gels, transferred to nitrocellulose filters and hybridized according to standard procedures with probes D, E and F (Figure 2C). Final washes were at 55°C in 0.2 × SSC, 0.1% SDS. All probes were synthesized in the presence of [α-^{32}P]dATP by primer extension using either the M13 universal primer (probe D) or oligo 3 (probe E in Figure 2C, see also Figure 1) or oligo 1 (probe F in Figure 2C, see also Figure 1) and a single stranded M13mp19 DNA containing the *Hin*dIII−*Pst*I gene fragment (Figure 2C). Probes D, E and F were isolated as described above after digestion with *Bgl*I (probe D and E) or *Nsi*I (probe F).

Construction of vectors expressing the progesterone receptor and its mutants in vivo

The cPR expression vectors were constructed using a genomic DNA fragment containing the *Hin*dIII−*Pst*I region of the cPR gene (see Figure 2C) and several overlapping cDNA fragments. These fragments were inserted into the *Eco*RI site of the eucaryotic expression vector $pKCR_2$ (Breathnach and Harris, 1983). The cPR expression vector series cPR0 to cPR6 was constructed (see Figure 3) using site-directed mutagenesis (Kumar *et al.*, 1986) and synthetic oligonucleotide adaptors. cPR0 contains the cDNA sequence between +29 and +2921 (see Figure 1), cPR1 the cDNA sequence between +367 and +2921 and the sequence 5'-GATCCTCGAGCCACC-3' upstream of it to provide an initiation codon within a Kozak consensus sequence. cPR4 was derived from cPR0 by deleting the sequences downstream from +1983 (amino acid 539) and adding at this position the sequence 5'-GAATTCACTCCTCAGGTGCAGGCTGCCTATCAGAAG-GTGGTGGCTGGTGTGGCCAATGCCCTGGCTCACAAATACCACTGA-3'. cPR5 was derived from cPR1 by deleting the sequences downstream from +1855 (amino acid 496) and adding at this position the sequence 5'-GATGTACAT-ATTGTGAGTTCCTGA-3'. cPR6 is identical to cPR1 up to +1718 (amino acid 451) and has the sequence 5'-TACTATATCATACCATTGTTCCTAA-3' downstream of it.

Construction of expression vectors for in vitro transcription/translation

For *in vitro* transcription the cPR sequences of cPR1 and cPR2 were inserted downstream of the T7 promoter in pGEM-1 (Promega), to yield cPR1G and cPR2G respectively. cPR0G is equivalent to cPR0 in pGEM-1 but it does not contain the sequences upstream of +314. *In vitro* transcription and *in vitro* translation was done essentially as described by Kumar (1986). The ^{35}S-labelled translation products were resolved on 7.5% SDS−polyacrylamide gels and visualized by fluorography using EN^{3}HANCE (NEN). For blocking the initiation of translation identical aliquots containing ~100 ng of *in vitro* synthesized RNA were incubated with increasing amounts of oligonucleotides anti-AUG 1 (5'-CTTCAC-CTCGGTCATGCTGC-3'), anti-AUG 128 (5'-GGGGCCGCGGCTCATCGG-GG-3') or with an unrelated oligonucleotide (5'-TTCACCCTCTCTTTCAGA-3') in a volume of 5 μl for 15 min at 70°C, before adding the translation mixture.

Determination of the transcriptional activation of the MMTV-LTR progesterone responsive element by quantitative S1 nuclease analysis and CAT assay

A reporter recombinant suitable for RNA determination by quantitative S1 nuclease analysis and containing the steroid hormone responsive element(s) of the MMTV LTR was constructed as follows. The MMTV LTR *Hae*III fragment (−635 to +122) was inserted upstream from the promoterless β-globin gene of pA0 (Zenke *et al.*, 1986) in which the SV40 early promoter sequences were replaced by a *Sma*I site-containing polylinker, yielding MMTV-globin (MG). Transfection experiments were carried out using HeLa or Cos-1 cells and the calcium phosphate technique (Banerji *et al.*, 1983). One μg of one of the cPR-expression vectors was transfected together with 2 μg of MG and 0.5 μg of pG1B (pGB in Sassone-Corsi *et al.*, 1985) as an internal control and 10 μg of Bluescribe plasmid as carrier ($3-5 \times 10^6$ cells/9 cm Petri dish). The cells were grown in the presence or absence of 5 nM R5020 as indicated in the figure legends. Cytoplasmic RNA was prepared 40 h after transfection by lysing the cells with 0.5% Nonidet P-40 as described (Zenke *et al.*, 1986). 15−25 μg of RNA were hybridized with an excess of a single stranded ^{32}P-5'-end-labelled DNA probe (see below) at 42°C in 50% (v/v) formamide, 400 mM NaCl, 40 mM Pipes, pH 6.5. After S1 nuclease digestion (140 units, 25°C, 2 h) the protected DNA sequences were resolved in 6% polyacrylamide/8.3 M urea gels. The single stranded DNA probe was prepared by first hybridizing a synthetic ^{32}P-5'-end-labelled oligonucleotide complementary to positions +39 to +60 of the rabbit β-globin gene (Zenke *et al.*, 1986) to a single stranded M13mp18 (M13MG) in which the *Sac*I−*Bam*HI fragment of MG containing the MMTV startsite of transcription has been inserted. After primer extension and digestion with *Sac*I, the single stranded DNA was purified by polyacrylamide gel electrophoresis and recovered by electroelution (Zenke *et al.*, 1986). CAT-assays were performed with approximately 20 μg of cytosolic protein prepared from transiently transfected HeLa cells [1 μg of expression vector, 1 μg MMTV-CAT (Cato *et al.*, 1986), 3 μg of the internal control recombinant pCH110 (Pharmacia), 5 μg Bluescribe plasmid carrier/$3-5 \times 10^6$ cells/9 cm Petri dish] as described (Gorman *et al.*, 1982) with the exception that the β-galactosidase activity encoded in the co-transfected pCH110 was first determined as described by the supplier in order to use 'normalized' amounts of cytosolic extracts.

Hormone binding analyses

All manipulations were done below 4°C. Cytosol of transfected HeLa or Cos-1 cells grown in the absence of steroid hormones was prepared by disrupting the cells in a glass-teflon motor-driven homogenizer (120 strokes) in TEBGN50-medium (Gronemeyer *et al.*, 1985) supplemented with 20 mM sodiummolybdate and 0.3 mM PMSF. After low speed (15 min, 1500 *g*) and high speed (60 min, 105 000 *g*) centrifugation, the cytosol was used directly for Scatchard analysis or incubation with 20 nM [^{3}H]R5020 (2 h, 0°C) followed by charcoal treatment and photoaffinity-labelling as described previously (Gronemeyer *et al.*, 1983; 1985). Non-treated cytosol (for immunoblotting) or the irradiated cytosol (for crosslinking analysis) was concentrated by precipitation in 50% ammonium sulphate followed by dialysis against Laemmli-sample buffer (Laemmli, 1970). Immunoblotting was carried out as described (Jeltsch *et al.*, 1986) using antiserum directed against the chicken progesterone receptor form B (Tuohimaa *et al.*, 1983), and analysis of the [^{3}H]R5020 crosslinked products was done by fluorography of 7.5% SDS−polyacrylamide gels using EN^{3}HANCE (NEN).

To analyse for 'tight' nuclear binding, cells incubated for 2 h at 37°C with 5 nM [^{3}H]R5020 ± 100-fold excess of non-radioactive hormone were disrupted as described above, except that the buffer was 340 mM sucrose, 2 mM Na_2 EDTA and 20 mM Tris−HCl, pH 7.4. After low speed centrifugation, the crude nuclear pellet was resuspended in the same buffer containing 150 mM NaCl, recentrifuged, and the pellet was extracted in the presence of 500 mM NaCl. After high speed centrifugation the radioactivity of aliquots of the nuclear extract and of the supernatants of the previous centrifugations was determined by liquid scintillation counting.

Acknowledgements

We are grateful to A.Staub and F.Ruffenach for the synthesis of oligonucleotide, to F.Jacob for preliminary sequencing work and to J.Eul for the purification and crosslinking of laying hen oviduct PR. The anti-cPR antiserum was a kind gift of P.Tuohimaa (Tampere) and the reporter plasmid MMTV-CAT was kindly provided by H.Ponta (Karlsruhe). We thank our colleagues, in particular S.Green and V.Kumar for helpful advice and discussion, the cell culture service for providing HeLa and Cos-1 cells and the secretarial staff, C.Werlé, B.Boulay, A.Landmann for their help in preparing this manuscript. B.T. is a post-doctoral fellow of the Medical Research Council of Canada. This research was supported by the CNRS (A.I.V.), the INSERM (grants CNAMTS), the Ministère de la Recherche et de la Technologie, the Fondation pour la Recherche Médicale and the Association pour la Recherche sur le Cancer.

References

Ahe,D.v.d., Renoir,J.-M., Buchou,T., Baulieu,E.-E. and Beato,M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2817−2821.

Arriza,J.L., Weinberger,C., Cerelli,G., Glaser,T.M., Handelin,B.L., Housman,D.E. and Evans,R.M. (1987) *Science*, **237**, 268−275.

Auffray,C. and Rougeon,F. (1980) *Eur. J. Biochem.*, **107**, 303−314.

Banerji,J., Olson,L. and Schaffner,W. (1983) *Cell*, **33**, 729–740.
Birnbaumer,M., Bell,R.C., Schrader,W.T. and O'Malley,B.W. (1984) *J. Biol. Chem.*, **259**, 1091–1098.
Birnbaumer,M., Schrader,W.T. and O'Malley,B.W. (1983) *J. Biol. Chem.*, **258**, 7331–7337.
Breathnach,R. and Harris,B. (1983) *Nucleic Acids Res.*, **11**, 7119–7136.
Cato,A.C.B., Miksicek,R., Schütz,G., Arnemann,J. and Beato,M. (1986) *EMBO J.*, **5**, 2237–2240.
Cato,A.C.B., Henderson,D. and Ponta,H. (1987) *EMBO J.*, **6**, 363–368.
Conneely,O.M., Sullivan,W.P., Toft,D.O., Birnbaumer,M., Cook,R.G., Maxwell,B.L., Zarucki-Schultz,T., Greene,G.L., Schrader,W.T. and O'Malley,B.W. (1986) *Science*, **233**, 767–770.
Danielsen,M., Northrop,J.P. and Ringold,G.M. (1986) *EMBO J.*, **5**, 2513–2522.
Dingwall,C., Dilworth,S.M., Black,S.J., Kearsey,S.E., Cox,L.S. and Laskey,R.A. (1987) *EMBO J.*, **6**, 69–74.
Earnshaw,W.C., Sullivan,K.F., Machlin,P.S., Cooke,C.A., Kaiser,D.A., Pollard,T.D., Rothfield,N.F. and Cleveland,D.W. (1987) *J. Cell Biol.*, **104**, 817–829.
Giguère,V., Hollenberg,S.M., Rosenfeld,M.G. and Evans,R.M. (1986) *Cell*, **46**, 645–652.
Godowski,P.J., Rusconi,S., Miesfeld,R. and Yamamoto,K.R. (1987) *Nature*, **325**, 365–368.
Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
Green,S. and Chambon,P. (1986) *Nature*, **324**, 615–617.
Green,S. and Chambon,P. (1987) *Nature*, **325**, 75–78.
Green,S., Gronemeyer,H. and Chambon,P. (1987) In Sluyser,M. (ed.), *Oncogenes and Hormones in Breast Cancer*. E.Horwood Ltd., Publishers, in press.
Gronemeyer,H. and Govindan,M.V. (1986) *Mol. Cell. Endocrinol.*, **46**, 1–19.
Gronemeyer,H., Govindan,M.V. and Chambon,P. (1985) *J. Biol. Chem.*, **260**, 6916–6925.
Gronemeyer,H., Green,S., Kumar,V., Jeltsch,J.M. and Chambon,P. (1987) In Sheridan,P.J., Blum,K. and Trachtenberg,M.C. (eds), *Steroid Receptors and Disorders: Cancer, Bone and Circulatory Disorders*. Marcel Dekker Inc., in press.
Gronemeyer,H., Harry,P. and Chambon,P. (1983) *FEBS Lett.*, **156**, 287–292.
Hollenberg,S.M., Giguère,V., Segui,P. and Evans,R.M. (1987) *Cell*, **49**, 39–46.
Horwitz,K.B., Wei,L.L., Sedlacek,S.M. and D'Arville,C.N. (1985) *Recent Prog. Horm. Res.*, **41**, 249–316.
Jeltsch,J.M., Krozowski,Z., Quirin-Stricker,C., Gronemeyer,H., Simpson,R.J., Garnier,J.M., Krust,A., Jacob,F. and Chambon,P. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5424–5428.
Kozak,M. (1984a) *Nucleic Acids Res.*, **12**, 857–872.
Kozak,M. (1984b) *Nucleic Acids Res.*, **12**, 3873–3893.
Krust,A., Green,S., Argos,P., Kumar,V., Walter,P., Bornert,J.M. and Chambon,P. (1986) *EMBO J.*, **5**, 891–897.
Kumar,V., Green,S., Staub,A. and Chambon,P. (1986) *EMBO J.*, **5**, 2231–2236.
Kumar,V., Green,S., Berry,M., Stack,G. and Chambon,P. (1987) *UCLA Symposium on Steroid Homone Action*, in press.
Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
Laughon,A., Carroll,S.B., Storfer,F.A., Riley,P.D. and Scott,M.P. (1985) *Cold Spring Harbor Symp. Quant. Biol.*, **50**, 253–262.
Logeat,F., Pamphile,R., Loosfelt,H., Jolivet,A., Fournier,A. and Milgrom,E. (1985) *Biochemistry*, **24**, 1029–1035.
Loosfelt,H., Atger,M., Misrahi,M., Guiochon-Mantel,A., Meriel,C., Logeat,F., Benarous,R. and Milgrom,E. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9045–9049.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560.
Miesfeld,R., Godowski,P.J., Maler,B.A. and Yamamoto,K.R. (1987) *Science*, **236**, 423–427.
Misrahi,M., Atger,M., d'Auriol,L., Loosfelt,H., Meriel,C., Fridlansky,F., Guiochon-Mantel,A., Galibert,F. and Milgrom,E. (1987) *Biochem. Biophys. Res. Commun.*, **143**, 740–748.
Mizusawa,S., Nishimura,S. and Seela,F. (1986) *Nucleic Acids Res.*, **14**, 1319–1324.
Rusconi,S. and Yamamoto,K.R. (1987) *EMBO J.*, **6**, 1309–1315.
Sassone-Corsi,P., Duboule,D. and Chambon,P. (1985) *Cold Spring Harbor Symp. Quant. Biol.*, **50**, 747–752.
Schrader,W.T., Birnbaumer,M.E., Hughes,M.R., Weigel,N.L., Grody,W.W. and O'Malley,B.W. (1980) *Recent Prog. Horm. Res.*, **37**, 583–633.
Simpson,R.J., Grego,B. and Gronemeyer,H. (1987) *Mol. Cell. Endocrinol.*, in press.
Sullivan,W.P., Beito,T.G., Proper,J., Krco,C.J. and Toft,D.O. (1986) *Endocrinology*, **119**, 1549–1557.
Tuohimaa,P., Renoir,J.-M., Radanyi,C., Mester,J., Joab,I., Buchou,T. and Baulieu,E.-E. (1984) *Biochem. Biophys. Res. Commun.*, **119**, 433–439.
Wharton,K.A., Yedvobnick,B., Finnerty,V.G. and Artavanis-Tsakonas,S. (1985) *Cell*, **40**, 55–62.
Zenke,M., Grundström,T., Matthes,H., Wintzerith,M., Schatz,C., Wildeman,A. and Chambon,P. (1986) *EMBO J.*, **5**, 387–397.

Received on September 11, 1987